

EUROPEAN PATENT APPLICATION

Application number: 88309194.4

Int. Cl. C12N 15/00 , C12P 21/02 ,  
A61K 37/02 , //G01N33/566

Date of filing 03.10.88

Priority 02.10.87 US 104329  
28.09.88 US 250785

Date of publication of application:  
03.05.89 Bulletin 89/18

Designated Contracting States:  
ES GR

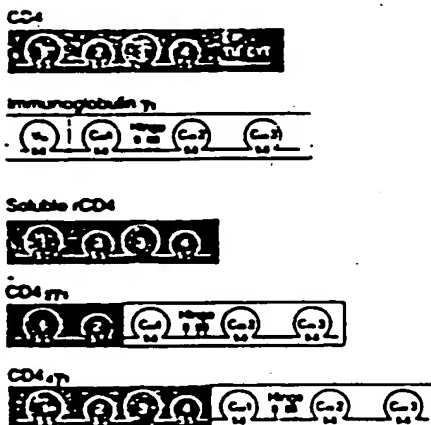
Applicant: GENENTECH, INC.  
460 Point San Bruno Boulevard  
South San Francisco California 94080(US)

Inventor: Capon, Daniel J.  
817 Oregon Avenue  
San Mateo California 94402(US)  
Inventor: Gregory, Timothy J.  
414 Pinehill Road  
Hillsborough California 94010(US)

Representative: Armitage, Ian Michael et al  
MEWBURN ELLIS & CO. 2/3 Cursitor Street  
London EC4A 1BO(GB)

Adhesion variants, nucleic acid encoding them and compositions comprising them.

Novel derivatives of cell surface proteins which are homologous to the immunoglobulin superfamily (adhesions) are provided. Amino acid sequence variations are introduced into adhesion, the most noteworthy of which are those in which the transmembrane and, preferably, cytoplasmic domains are rendered functionally inactive, and in which adhesion extracellular domains replace an immunoglobulin variable region. These variants are useful in therapy or diagnostics, in particular, CD4 variants are therapeutically useful in the treatment of HIV infections



89131855

1 6 4 5

EP 0 314 317 A1

BEST AVAILABLE COPY

Applicants: Norbert Schulke et al.  
Serial No.: 10/804,802  
Filed : March 19, 2004  
Exhibit 10

## ADHESION VARIANTS, NUCLEIC ACID ENCODING THEM AND COMPOSITIONS COMPRISING THEM

This application relates to compositions for antiviral or immunomodulatory therapy, in particular, it relates to compositions useful in the treatment of Human Immunodeficiency Virus (HIV) infections.

The primary immunologic abnormality resulting from infection by HIV is the progressive depletion and functional impairment of T lymphocytes expressing the CD4 cell surface glycoprotein (H. Lane et al. *Ann. Rev. Immunol.* 3:477 [1985]). CD4 is a non-polymorphic glycoprotein with homology to the immunoglobulin gene superfamily (P. Maddon et al. *Cell* 42:93 [1985]). Together with the CD8 surface antigen, CD4 defines two distinct subsets of mature peripheral T cells (E. Reinherz et al. *Cell* 19:821 [1980]), which are distinguished by their ability to interact with nominal antigen targets in the context of class I and class II major histocompatibility complex (MHC) antigens, respectively (S. Swain, *Proc. Natl. Acad. Sci.* 78:7101 [1981]; E. Engleman et al. *J. Immunol.* 127:2124 [1981]; H. Spitz et al. *J. Immunol.* 129:1563 [1982]; W. Biddison et al. *J. Exp. Med.* 156:1065 [1982], and D. Wilde et al. *J. Immunol.* 131:2178 [1983]). For the most part, CD4<sup>+</sup> T cells display the helper inducer T cell phenotype (E. Reinherz, *supra*), although CD4<sup>+</sup> T cells characterized as cytotoxic suppressor T cells have also been identified (Y. Thomas et al. *J. Exp. Med.* 154:459 [1981]; S. Meuer et al. *Proc. Natl. Acad. Sci. USA* 79:4395 [1982]; and A. Krensky et al. *Proc. Natl. Acad. Sci. USA* 79:2365 [1982]). The loss of CD4 helper inducer T cell function probably underlies the profound defects in cellular and humoral immunity leading to the opportunistic infections and malignancies characteristic of the acquired immunodeficiency syndrome (AIDS) (H. Lane *supra*).

Studies of HIV-1 infection of fractionated CD4 and CD8 T cells from normal donors and AIDS patients have revealed that depletion of CD4<sup>+</sup> T cells results from the ability of HIV-1 to selectively infect, replicate in, and ultimately destroy this T lymphocyte subset (D. Klatzmann et al. *Science* 225:59 [1984]). The possibility that CD4 itself is an essential component of the cellular receptor for HIV-1 was first indicated by the observation that monoclonal antibodies directed against CD4 block HIV-1 infection and syncytia induction (A. Daigleish et al. *Nature [London]* 312:767 [1984]; J. McDougal et al. *J. Immunol.* 135:3151 [1985]). This hypothesis has been confirmed by the demonstration that a molecular complex forms between CD4 and gp120, the major envelope glycoprotein of HIV-1 (J. McDougal et al. *Science* 231:382 [1986]; and the finding that HIV-1 tropism can be conferred upon ordinarily non-permissive human cells following the stable expression of a CD4 cDNA (P. Maddon et al. *Cell* 47:333 [1986]). Furthermore, the neurotropic properties of HIV-1, reflected by a high incidence of central nervous system dysfunction in HIV-1 infected individuals (W. Snider et al. *Ann. Neurol.* 14:403 [1983]), and the ability to detect HIV-1 in the brain tissue and cerebrospinal fluid of AIDS patients (G. Shaw et al. *Science* 227:177 [1985]; L. Epstein *AIDS Res.* 1:447 [1985]; S. Koenig, *Science* 233:1089 [1986]; D. Ho et al. *N. Engl. J. Med.* 313:1498 [1985]; J. Levy et al. *Lancet* ii:586 [1985]) appears to have its explanation in the expression of CD4 in cells of neuronal, glial and monocyte/macrophage origin (P. Maddon, *Cell* 47:444 [1986]; I. Funke et al. *J. Exp. Med.* 165:1230 [1986]; B. Tournelle et al. *Science* 234:610 [1986]).

In addition to determining the susceptibility to HIV-1 infection, the manifestation of cytopathic effects in the infected host cell appears to involve CD4. Antibody to CD4 was found to inhibit the fusion of uninfected CD4<sup>+</sup> T cells with HIV-1 infected cells *in vitro*; moreover, the giant multinucleated cells produced by this event die shortly after being formed resulting in the depletion of the population of CD4<sup>+</sup> cells (J. Lifson et al. *Science* 232:1123 [1986]). Formation of syncytia also requires gp120 expression, and can be elicited by coculturing CD4-positive cell lines with cell lines expressing the HIV-1 env gene in the absence of other viral structural or regulatory proteins (J. Sodroski et al. *Nature* 322:470 [1986]; J. Lifson et al. *Nature* 323:725 [1986]). Thus, in mediating both the initial infection by HIV-1 as well as eventual cell death, the interaction between gp120 and CD4 constitutes one of several critical entry points in the viral life cycle amenable to therapeutic intervention (H. Mitsuya et al. *Nature* 325:773 [1987]).

The known sequence of the CD4 precursor predicts a hydrophobic signal peptide an extracellular region of approximately 370 amino acids, a highly hydrophobic stretch with significant identity to the membrane-spanning domain of the class II MHC beta chain, and a highly charged intracellular sequence of 40 residues (P. Maddon, *Cell* 42:93 [1985]). The extracellular domain of CD4 consists of four contiguous regions each having amino acid and structural similarity to the variable and joining (V-J) domains of immunoglobulin light chains as well as related regions in other members of the immunoglobulin gene superfamily (a subclass of which are defined herein by the coined term "adhesions". These structurally similar regions of CD4 are termed the V<sub>1</sub>J<sub>1</sub>, V<sub>2</sub>J<sub>2</sub>, V<sub>3</sub>J<sub>3</sub> and V<sub>4</sub>J<sub>4</sub> domains (denominated 1-4 in Fig. 3).

A successful strategy in the development of drugs for the treatment of many receptor mediated abnormalities has been the identification of antagonists which block binding of the natural ligand. Since the

CD4 adhesion ordinarily binds to the recognition sites of the HIV envelope it would appear to be a candidate for therapeutically sequestering these HIV sites, thereby blocking viral infectivity. However, full length CD4 and other adhesions are cell membrane proteins which are anchored in the lipid bilayer of lymphocytes. The presence of membrane components will be undesirable from the standpoint of manufacturing and purification. In addition, since adhesions are normally present only on cell surfaces, it would be desirable to produce adhesions in a form which is more stable in the circulation. Additionally, even truncated, soluble CD4 adhesion (generally referred to as CD4T) may not be optimally effective as a therapeutic since it possesses a relatively short biological half-life, binds to HIV no better than cell surface CD4, may not cross the placental or other biological barriers and since it merely sequesters the HIV recognition sites without in itself bearing an infected-cell killing or virus killing functionality.

Accordingly, it is desirable to produce soluble, secreted adhesions. It is also desirable to produce CD4 derivatives useful in the treatment of AIDS and related conditions, in a manner essentially unaffected by the extreme degree of genetic variation observed among various HIV-1 isolates and their respective env polypeptides (J. Coffin, Cell 46:1 [1986]). It is further desirable to prepare adhesions fused to other polypeptides in order to provide molecules with novel functionalities such as those described above for therapeutic use, or diagnostic reagents for the *in vitro* assay of adhesions or their ligands. In particular, it is desirable to prepare molecules for directing toxins or effector molecules (for example the Fc domain of immunoglobulin) to cells bearing receptors for the adhesions, e.g. HIV gp120 in the case of CD4 and for use in facilitating purification of the adhesions. It is further desirable to provide stable, highly purified adhesion preparations.

The present invention provides nucleic acid encoding an amino acid sequence variant of an adhesion, in particular a variant in which the trans-membrane domain is modified so that it is no longer capable of becoming lodged in the cell membrane, i.e. inactivated, for example, by deletion or substitution. In the case of CD4 such variants are termed CD4

Variant adhesions are produced by a method comprising (a) transforming a host cell with nucleic acid encoding an amino acid sequence variant of an adhesion, (b) culturing the host cell and (c) recovering the variant adhesion from the host cell culture media.

In another embodiment the invention provides an adhesion variant selected from the group consisting of (a) an adhesion amino acid sequence variant having an inactivated transmembrane domain and (b) a polypeptide comprising an adhesion extracellular domain fused to the sequence of a polypeptide which is different from the adhesion this latter selected from a cytotoxin, an immunogen or an immunoglobulin constant domain.

In a preferred embodiment a polypeptide comprising a gp120 binding domain of the CD4 adhesion is fused at its C-terminus to an immunoglobulin constant domain, or is linked to a cytotoxic polypeptide such as ricin.

The CD4 adhesion variants provided herein are purified and formulated in pharmacologically acceptable vehicles for administration to patients in need of antiviral or immunomodulatory therapy, in particular patients infected with HIV.

#### Brief Description of the Drawings

Figs. 1a-1c depict the amino acid and nucleotide sequence of a secreted form of the CD4 adhesion (CD4T). Other forms of CD4T terminate at Ser 366 or Pro 368. The signal processing site is designated with an arrow.

Figs. 2a-2c depict the amino acid and nucleotide sequence of a fusion of the herpes gD leader and N-terminal 27 residues to the putative mature N-terminus of CD4T.

Fig. 3 depicts the structural elements of the native and soluble CD4 antigen, the native immunoglobulin G heavy ( $\gamma$ ) chain and two exemplary heavy chain-CD4 chimeras.

Figs. 4a-4b are a map of the linked immunoglobulin  $\gamma$  chain fragment employed in the preparation of CD4 fusions. Insert sites are designated  $\gamma$ 1 and Fc.

Fig. 5 is a map of a human light chain fragment useful for CD4 fusions at the arrow flanked by V<sub>L</sub>, (light variable and joining) and C<sub>L</sub> (light constant).

#### Detailed Description

Adhesions are cell surface polypeptides having an extracellular domain which is homologous to a member of the immunoglobulin gene superfamily, excluding, however, highly polymorphic members of this superfamily selected from the group of class I and class II major histocompatibility antigens, immunoglobulins and T-cell receptor  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  chains. Examples of adhesions include CD2, CD4, CD8, CD28, the  $\gamma$ ,  $\delta$  and  $\epsilon$  chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated glycoprotein (MAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen. Homologous as defined herein means having the sequence of a member of the immunoglobulin gene superfamily or having a sequence therewithin which has substantially the same as (or a greater degree of) amino acid sequence homology to a known member of the superfamily as the specific examples given above have to the sequence of an immunoglobulin variable or constant domain. Preferred adhesions are CD4, CD8 and high affinity IgE receptor.

This invention is particularly concerned with amino acid sequence variants of adhesions. Amino acid sequence variants of adhesions are prepared with various objectives in mind, including increasing the affinity of the adhesion for its binding partner, facilitating the stability, purification and preparation of the adhesion, increasing its plasma half life, improving therapeutic efficacy as described above in the background, introducing additional functionalities and lessening the severity or occurrence of side effects during therapeutic use of the adhesion. Amino acid sequence variants of adhesions fall into one or a combination of the following classes: insertional, substitutional or deletional variants.

Insertional amino acid sequence variants are those in which one or more amino acid residues extraneous to the adhesion are introduced into a predetermined site in the adhesion including the C or N terminus. Such variants are referred to as fusions of the adhesion and a different polypeptide are produced. Such other polypeptides contain sequences other than those which are normally found in the adhesion at the inserted position. Several groups of fusions are contemplated herein. Immunologically active adhesion fusions comprise an adhesion and a polypeptide containing a non-adhesion epitope. The non-adhesion epitope is any immunologically competent polypeptide, i.e., any polypeptide which is capable of eliciting an immune response in the animal to which the fusion is to be administered or which is capable of being bound by an antibody raised against the non-adhesion polypeptide. Typical non-adhesion epitopes will be those which are borne by allergens, autoimmune epitopes, or other potent immunogens or antigens recognized by pre-existing antibodies in the fusion recipient, including bacterial polypeptides such as trpLE, beta-galactosidase, viral polypeptides such as herpes gD protein, and the like. Immunogenic fusions are produced by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding an immunogenic polypeptide. It is preferable that the immunogenic fusion be one in which the immunogenic sequence is joined to or inserted into the adhesion antigen or fragment thereof by a peptide bond(s). These products therefore consist of a linear polypeptide chain containing adhesion epitopes and at least one epitope foreign to the adhesion. It will be understood that it is within the scope of this invention to introduce the epitopes anywhere within the adhesion molecule or fragment thereof. Such fusions are conveniently made in recombinant host cells or by the use of bifunctional cross-linking agents. The use of a cross-linking agent to fuse the adhesion to the immunogenic polypeptide is not as desirable as a linear fusion because the cross-linked products are not as easily synthesized in structurally homogeneous form.

These immunogenic insertions are particularly useful when formulated into a pharmacologically acceptable carrier and administered to a subject in order to raise antibodies against the adhesion, which antibodies in turn are useful in diagnostics or in purification of adhesion by immunoaffinity techniques known per se. Alternatively, in the purification of adhesions, binding partners for the fused non-adhesion polypeptide, e.g., antibodies, receptors or ligands, are used to adsorb the fusion from impure admixtures, after which the fusion is eluted and, if desired, the adhesion is recovered from the fusion, e.g. by enzymatic cleavage.

Other fusions, which may or may not also be immunologically active, include fusions of the adhesion sequence with a signal sequence heterologous to the adhesion, fusions of transmembrane-modified CD4 adhesions, for example, to polypeptides having enhanced plasma half life (ordinarily > about 20 hours) such as immunoglobulin chains or fragments thereof, and fusions with cytotoxic functionalities. Signal sequence fusions are employed in order to more expeditiously direct the secretion of the adhesion. The heterologous signal replaces the native adhesion signal, and when the resulting fusion is recognized, i.e. processed and cleaved by the host cell, the adhesion is secreted. Signals are selected based on the intended host cell and may include bacterial, yeast, mammalian and viral sequences. The herpes gD glycoprotein signal is suitable for use in mammalian expression systems.

Plasma proteins which have enhanced plasma half-life longer than that of transmembrane modified CD4 include serum albumin, immunoglobulins, apolipoproteins, and transferrin. Preferably the adhesion-plasma

protein fusion is not significantly immunogenic in the animal in which it is used and the plasma protein does not cause undesirable side effects in patients by virtue of its normal biological activity

In a specific embodiment the adhesion constant and/or variable region-like domain is conjugated with an immunoglobulin constant region sequence. The resulting products are referred to herein as immunoadhesions. Immunoglobulins and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Patent 4,745,055; EP 256,654; Faulkner et al., *Nature* 298:286 (1982); EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 (1979); Köhler et al., *P.N.A.S. USA* 77:2197 (1980); Raso et al., *Cancer Res.* 41:2073 (1981); Morrison et al., *Ann. Rev. Immunol.* 2:239 (1984); Morrison, *Science* 229:1202 (1985); Morrison et al., *P.N.A.S. USA* 81:6851 (1984); EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. patent 4,444,878; WO 88/03565; and EP 68,763 and references cited therein.

Ordinarily, the domains of adhesions that are homologous to immunoglobulins and extracellular in their native environment are fused C-terminally to the N-terminus of the constant region of immunoglobulins in place of the variable region(s)-like thereof, retaining at least functionally active hinge, CH2 and CH3 domains of the constant region of an immunoglobulin heavy chain. This ordinarily is accomplished by constructing the appropriate DNA sequence and expressing it in recombinant cell culture. Immunoglobulins and other polypeptides having enhanced plasma half life are fused to the extracellular or ligand binding domains of other adhesions in the same fashion.

The boundary domains for the CD4 V-like regions (V1-V4) are, respectively, about 100-109, about 175-184, about 289-298, and about 360-369 (based on the precursor CD4 amino acid sequence in which the initiating met is -25, Fig. 1a). CD4 sequences containing any of the CD4 V domains are fused to the immunoglobulin sequence. It is preferable that the V1V2 or V1V2V3V4 be fused at their C-termini to the immunoglobulin constant region. The precise site at which the fusion is made is not critical; the boundary domains noted herein are for guidance only and other sites neighboring or within the V regions may be selected in order to optimize the secretion or binding characteristics of the CD4. The optimal site will be determined by routine experimentation. In general, it has been found that the fusions are expressed intracellularly, but a great deal of variation is encountered in the degree of secretion of the fusions from recombinant hosts. For instance, the following table demonstrates the various immunoglobulin fusions that have been obtained by the method of this invention. In all examples of CD4 immunoadhesions, the CD4 signal was used to direct secretion from 293 cells. Lower case m represents murine origin, while the lower case h designates human origin. V and C are abbreviations for immunoglobulin variable and constant domains respectively. The numerical subscripts indicate the number of parenthetical units found in the designated multimer. It will be understood that the chains of the multimers are believed to be disulfide bonded in the same fashion as native immunoglobulins. The CD4 immunoadhesions typically contained either the first N-terminal 366 residues of CD4 (CD4<sub>1</sub>) or the first 180 N-terminal residues of CD4 (CD4<sub>2</sub>) linked at their C-terminus to the α (light) chain or IgG1 heavy chain constant region (γ1).

89131855

1649

Table I

Transfected Gene	Secreted Product
$mV_{\kappa}C_{\kappa}$	$mV_{\kappa}C_{\kappa}$ and/or $(mV_{\kappa}C_{\kappa})_2$
$mV_{\gamma 1}C_{\gamma 1}$	ND
$mV_{\kappa}C_{\kappa} + mV_{\gamma 1}C_{\gamma 1}$	$(mV_{\kappa}C_{\kappa})_2(mV_{\gamma 1}C_{\gamma 1})_2 +$ $mV_{\kappa}C_{\kappa}$ and/or $(mV_{\kappa}C_{\kappa})_2$
$hCD4-mC_{\kappa}$	$hCD4-mC_{\kappa}$ and/or $(hCD4-mC_{\kappa})_2$
$hCD4-mC_{\gamma 1}$	ND
$hCD4-mC_{\kappa} + hCD4-mC_{\gamma 1}$	$(hCD4-mC_{\kappa})_2(hCD4-mC_{\gamma 1})_2 +$ $hCD4-mC_{\kappa}$ and/or $(hCD4-mC_{\kappa})_2$
$hCD4-hC_{\kappa}$	$hCD4-hC_{\kappa}$ and/or $(hCD4-hC_{\kappa})_2$
$hCD4-hC_{\gamma 1}$	$(hCD4-hC_{\gamma 1})_2$
$hCD4-hC_{\kappa} + hCD4-hC_{\gamma 1}$	$(hCD4-hC_{\kappa})_2(hCD4-hC_{\gamma 1})_2 +$ $hCD4-hC_{\kappa}$ and/or $(hCD4-hC_{\kappa})_2$
$mV_{\kappa}C_{\kappa} + hCD4-hC_{\gamma 1}$	$(mV_{\kappa}C_{\kappa})_2(hCD4-hC_{\gamma 1})_2 +$ $mV_{\kappa}C_{\kappa}$ and/or $(mV_{\kappa}C_{\kappa})_2$

\*ND - Not detected

It is interesting to observe from this table that the CD4-human heavy chain immunoadheson was secreted as a dimer whereas the analogous murine construction was not detected (this not excluding the intracellular accumulation of the protein, however). The ability of the  $hCD4-hC_{\gamma 1}$  transformants to produce heavy chain dimer was unexpected since previous work had suggested that immunoglobulin heavy chains are not secreted unless the hosts are cotransformed with nucleic acid encoding both heavy and light chain (Valle et al., Nature 241:338 [1981]). According to this invention, CD4-IgG immunoadheson chimeras are readily secreted wherein the CD4 epitope is present in light chain dimers, heavy chain monomers or dimers, and heavy and light chain tetramers wherein the CD4 epitope is present fused to one or more light or heavy chains, including heterotetramers wherein up to and including all four variable region analogues are derived from CD4. Where light-heavy chain non-CD4 variable domain is present, a heterofunctional antibody thus is provided.

Various exemplary hetero- and chimeric immunoadheson antibodies produced in accordance with this invention are schematically diagrammed below. "A" means at least a portion of the extracellular domain of an adhesion containing its ligand binding site;  $V_L$ ,  $V_H$ ,  $C_L$  and  $C_H$  represent light or heavy chain variable or constant domains of an immunoglobulin;  $n$  is an integer; and  $Y$  designates a covalent cross-linking moiety

(a)  $AC_L$ .

89131855

- (b)  $AC_L-AC_L$
- (c)  $AC_L-[AC_L-AC_L-AC_L-V_HC_H, V_LC_L-AC_L, \text{ or } V_LC_L-V_HC_H]$
- (d)  $AC_L-AC_L-[AC_L-AC_L-AC_L-V_HC_H, V_LC_L-AC_L, \text{ or } V_LC_L-V_HC_H]$
- (e)  $AC_L-V_HC_H-[AC_L-AC_L-AC_L-V_HC_H, V_LC_L-AC_L, \text{ or } V_LC_L-V_HC_H]$
- (f)  $V_LC_L-AC_L-[AC_L-AC_L-AC_L-V_HC_H, V_LC_L-AC_L, \text{ or } V_LC_L-V_HC_H]$

- (g)  $[A-Y]_n-[V_LC_L-V_HC_H]$

The structures shown in this table show only key features, e. g. they do not show joining (J) or other domains of the immunoglobulins, nor are disulfide bonds shown. These are omitted in the interests of brevity. However, where such domains are required for binding activity they shall be construed as being present in the ordinary locations which they occupy in the adhesion, immunoadhesion or immunoglobulin molecules as the case may be. These examples are representative of divalent antibodies, more complex structures would result by employing immunoglobulin heavy chain sequences from other classes e.g. IgM. The immunoglobulin  $V_LV_H$  antibody combining site also designated as the companion immunoglobulin preferably is capable of binding to a predetermined antigen.

Suitable companion immunoglobulin combining sites and fusion partners are obtained from IgG-1, -2, -3, or 4- subtypes, IgA, IgE, IgD or IgM, but preferably IgG-1.

A preferred embodiment is a fusion of an N-terminal portion of CD4, which contains the binding site for the gp120 envelope protein of HIV, to the C-terminal  $F_c$  portion of an antibody, containing the effector functions of immunoglobulin G. There are two preferred embodiments of this sort: in one, the entire heavy chain constant region is fused to a portion of CD4; in another, a sequence beginning in the hinge region just upstream of the papain cleavage site which defines IgG  $F_c$  chemically (residue 216, taking the first residue of heavy chain constant region to be 114 [Kabat et al., "Sequences of Proteins of Immunological Interest" 4th Ed., 1987], or analogous sites of other immunoglobulins) is fused to a portion of CD4. These embodiments are described in the examples.

More particularly, those variants in which one or more immunoglobulin-like domains of an adhesion are substituted for the variable region of an immunoglobulin chain are believed to exhibit improved *in vivo* plasma half life. These chimeras are constructed in a fashion similar to chimeric antibodies in which a variable domain from a antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023, Munro, Nature 312: (13 December 1984); Neuberger et al., Nature 312: (13 December 1984); Sharon et al., Nature 309: (24 May 1984); Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984); Morrison et al., Science 229:1202-1207 (1985); and Boulikas et al., Nature 312:643-646 (13 December 1984). The DNA encoding a CD4 V-like region is cleaved by a restriction enzyme at or proximal to a DNA encoding a boundary domain and at a point at or near the DNA encoding N-terminal 5' end of the mature antigen (where use of a non-CD4 lead is contemplated) or at or proximal to the 5' end of the N-terminal coding region for the antigen (where a native CD4 signal is employed). This DNA fragment then is readily inserted into DNA encoding an immunoglobulin light or heavy chain constant region and, if necessary, tailored by deletional mutagenesis. Preferably, this is a human immunoglobulin when the variant is intended for *in vivo* therapy for humans. DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See for example, Adams et al., Biochemistry 19:2711-2719 (1980); Gough et al., Biochemistry 19:2702-2710 (1980); Dolby et al., P.N.A.S. USA 77:6027-6031 (1980); Rice et al., P.N.A.S. USA 79:7862-7865 (1982); Falkner et al., Nature 298:286-288 (1982); and Morrison et al., Ann. Rev. Immunol. 2:239-256 (1984).

DNA encoding the immunoglobulin or immunoadhesion chimeric chain(s) is transfected into a host cell for expression. If the host cell is producing a heavy chain immunoglobulin prior to transfection then one need only transfect with the adhesion fused light chain to produce a heteroantibody. Similarly, if the host cell is expressing a light chain then DNA encoding a heavy chain adhesion fusion is transfected to produce a heteroantibody. The aforementioned immunoglobulins having one or more arms bearing the adhesion domain and one or more arms bearing companion variable regions result in dual specificity for adhesion ligand and for an antigen. These are produced by the above-described recombinant methods or by *in vitro* procedures. In the latter case, for example, Fab 12 fragments of the adhesion fusion and an immunoglobulin are prepared, the Fab 12 fragments converted to Fab fragments by reduction under mild reducing conditions and then reoxidized in each other's presence under acidic conditions in accord with methods known per se. See also U.S. patent 4,444,878.

In an alternative method for producing a heterofunctional antibody, host cells producing an adhesion-immunoglobulin fusion, e.g. transfected myelomas, also are fused with B cells or hybridomas which secrete antibody having the desired companion specificity for an antigen. Heterobifunctional antibody is recovered

89131855

from the culture medium of such hybridomas, and thus may be produced somewhat more conveniently than by conventional *in vitro* resorting methods (EP 68,763).

Another group of fusions are those in which an adhesion is conjugated with a toxic substance, e.g. a polypeptide such as ricin (including deglycosylated ricin A chain), diphtheria toxin A, or a non-peptidyl cytotoxin. Where the toxin is a polypeptide it is convenient to cross-link the polypeptide to the adhesion or its transmembrane-deleted variant by conventional *in vitro* protein cross-linking agents (for suitable methods for linking ricin A chain or deglycosylated A chain to CD4 see, for example, Duncan et al., "Analyt Biochem" 132:68-73 [1983]; Thorpe et al., "Cancer Res." 47:5924 [1987]; and Ghotie et al., "Cancer Res" 48:2610 [1988]) or by recombinant synthesis as a fusion (see for example, U.S. Patent 4,765,382). Alternatively, where companion antibodies are anti-ricin antibody immunoglobulin variable domains, such immunoglobulin heteroantibodies are employed to deliver ricin to HIV infected cells following the general procedure of Raso et al., Cancer Research, 41:2073 (1981).

Another class of adhesion variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from an adhesion sequence. Typically, the transmembrane and cytoplasmic domains of adhesions are deleted. In the case of CD4, at least residues 368 to 395 (the transmembrane region), and ordinarily 396-433 as well (the cytoplasmic domain), will be deleted to obtain secreted forms of this adhesion. Parenthetically, the amino acid residues follow the numbers given for mature CD4 as noted, for example, in figures 1a - 1c.

Substitutional variants are those in which at least one residue in the adhesion sequence has been removed and a different residue inserted in its place. The native N-terminal residue for mature CD4 is now known to be lysine. Thus, the sequence shown in Fig. 1, with an N-terminal asparagine, is an amino acid sequence variant of native mature CD4. Table 2 below describes substitutions which in general will result in fine modulation of the characteristics of the CD antigen.

TABLE 2

Original Residue	Exemplary Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser; ala
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in adhesion properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteinyl or prolyl is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a



residue having a bulky side chain e.g. phenylalanyl, is substituted for (or by) one not having a side chain, e.g. glycyl.

A preferred class of substitutional or deletional variants are those involving the transmembrane region of the adhesion. The transmembrane region of the adhesion is a highly hydrophobic or lipophilic domain that is the proper size to span the lipid bilayer of the cellular membrane. It is believed to anchor the adhesion in the cell membrane.

Deletion or substitution of the transmembrane domain will facilitate recovery and provide a soluble form of the adhesion by reducing its cellular or membrane lipid affinity and improving its water solubility. If the transmembrane and cytoplasmic domains are deleted one avoids the introduction of potentially immunogenic epitopes either by exposure of otherwise intracellular polypeptides that might be recognized by the body as foreign or by insertion of heterologous polypeptides that are potentially immunogenic. A principal advantage of the transmembrane deleted adhesion is that it is secreted into the culture medium of recombinant hosts. This variant is water soluble and does not have an appreciable affinity for cell membrane lipids, thus considerably simplifying its recovery from recombinant cell culture.

It will be amply apparent from the foregoing discussion that substitutions, deletions, insertions or any combination thereof are introduced to arrive at a final construct. As a general proposition, all variants will not have a functional transmembrane domain and preferably will not have a functional cytoplasmic sequence. This is generally accomplished by deletion of the relevant domain, although adequate insertional or substitutional mutagens also can be effective for this purpose. For example, the transmembrane domain is substituted by any amino acid sequence, e.g. a random or homopolynucleic sequence of about 5 to 50 serine, threonine, lysine, arginine, glutamine, aspartic acid and like hydrophilic residues, which altogether exhibit a hydrophilic hydropathy profile, so that it is secreted into the culture medium of recombinant hosts. This variant should also be considered to be an adhesion variant.

These variants ordinarily are prepared by site specific mutagenesis of nucleobases in the DNA encoding the adhesion, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. However, variant adhesions also are prepared by *in vitro* synthesis. Obviously, variations made in the DNA encoding the variant adhesions must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure deleterious to expression (EP 75,444A). The CD4 variants typically exhibit the same gp120 binding activity as does the naturally-occurring prototype, although variants also are selected in order to modify the characteristics of the CD4 adhesion as indicated above.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed adhesion variants screened for the optimal combination of desired activities. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis.

Adhesion variants that are not capable of binding HIV gp120 are useful nonetheless as immunogens for raising antibodies to the adhesion or as immunoassay kit components (labelled, as a competitive reagent for gp120 assay, or unlabelled as a standard for an adhesion assay) so long as at least one adhesion epitope remains active.

The DNA encoding adhesions is obtained by known procedures. See Reihertz et al. and Maccon et al. *op cit*. In general, prokaryotes are used for cloning of CD4 variant DNA sequences. For example, *E. coli* K12 strain 294 (ATCC No. 31446) is particularly useful. Other microbial strains which may be used include *E. coli* B, UM101 and *E. coli* 1776 (ATCC No. 31537). These examples are illustrative rather than limiting.

DNA encoding the variant adhesions are inserted for expression into vectors containing promoters and control sequences which are derived from species compatible with the intended host cell. The vector ordinarily, but need not, carry a replication site as well as one or more marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using a derivative of pBR322 which is a plasmid derived from an *E. coli* species (Bolivar, et al., *Gene* 2: 95 (1977)). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA constructions.

Promoters suitable for use with prokaryotic hosts illustratively include the  $\beta$ -lactamase and lactose promoter systems (Chang et al., *Nature* 275: 615 (1978); and Goeddel et al., *Nature* 281: 544 (1979)), alkaline phosphatase, the tryptophan (*trp*) promoter system (Goeddel, *Nucleic Acids Res.* 8: 4057 (1980); and EPO Appin Publ. No. 36,776) and hybrid promoters such as the *tac* promoter (H. de Boer et al., *Proc*

Natl. Acad. Sci. USA 80 21-25 (1983)) However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to DNA encoding the adhesion variant using linkers or adaptors to supply any required restriction sites (Siebenlist et al., Cell 20 269 (1980)). Promoters for use in bacterial systems also will contain a Shine-Dalgarno (SD) sequence operably linked to the DNA encoding the antigen.

In addition to prokaryotes eukaryotic microbes such as yeast cultures also are useful as cloning or expression hosts. *Saccharomyces cerevisiae*, or common baker's yeast is the most commonly used eukaryotic microorganism although a number of other strains are commonly available. For expression in *Saccharomyces*, the plasmid YRp7 for example, (Stinchcomb, et al., Nature 282: 39 (1979); Kingsman et al., Gene 7 141 (1979); Tschemper et al., Gene 10: 157 (1980)) is commonly used. This plasmid already contains the *trp1* gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC no. 4076 or PEP4-1 (Jones, Genetics 85: 12 (1977)). The presence of the *trp1* lesion as a characteristic of the yeast host cell genome then provides an effective means of selection by growth in the absence of tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem. 255 2073 (1980)) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg. 7 149 (1968), and Holland, Biochemistry 17, 4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytocrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., European Patent Publication No. 73 657A. Yeast enhancers also are advantageously used with yeast promoters.

Promoters controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example the genomes of viruses such as, polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters e.g. the beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiers et al., Nature, 273 113 (1978). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment, Greenaway, P.J. et al., Gene 18 355-360 (1982). Of course, promoters from the host cell or related species also are useful herein.

DNA transcription in higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually from about 10 to 300bp, that act to increase the transcription initiation capability of a promoter. Enhancers are relatively orientation and position independent having been found 5' (Lairins, L. et al., Proc. Natl. Acad. Sci. 78 993 (1981)) and 3' (Lusk, M.L., et al., Mol. Cell Bio. 3 1108 (1983)) to the transcription unit, within an intron (Banerji, J.L. et al., Cell 33: 729 (1983)) as well as within the coding sequence itself (Osborne, T.F., et al., Mol. Cell Bio. 4: 1293 (1984)). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding the adhesion.

Expression vector systems generally will contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented medium. Two examples are, CHO DHFR<sup>-</sup> cells and mouse LTK<sup>-</sup> cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway they cannot survive unless the missing nucleotides

are provided in a supplemented medium. An alternative to supplementing the medium is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non supplemented media.

The second category is dominant selection which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin, Southern P. and Berg, P. *J. Molec. Appl. Genet.* 1: 327 (1982), mycophenolic acid, Mulligan, R.C. and Berg, P. *Science* 209: 1422 (1980) or hygromycin, Sugden, B. et al., *Mol. Cell Biol.* 5: 410-413 (1985). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), gpt (mycophenolic acid) or hygromycin, respectively.

"Amplification" refers to the increase or replication of an isolated region within a cell's chromosomal DNA. Amplification is achieved using a selection agent e.g. methotrexate (MTX) which inactivates DHFR. Amplification or the making of successive copies of the DHFR gene results in greater amounts of DHFR being produced in the face of greater amounts of MTX. Amplification pressure is applied notwithstanding the presence of endogenous DHFR, by adding ever greater amounts of MTX to the media. Amplification of a desired gene can be achieved by cotransfecting a mammalian host cell with a plasmid having a DNA encoding a desired protein and the DHFR or amplification gene permitting cotransformation. One ensures that the cell requires more DHFR, which requirement is met by replication of the selection gene, by selecting only for cells that can grow in the presence of ever-greater MTX concentration. So long as the gene encoding a desired heterologous protein has cotransformed with the selection gene replication of this gene gives rise to replication of the gene encoding the desired protein. The result is that increased copies of the gene, i.e. an amplified gene encoding the desired heterologous protein express more of the desired heterologous protein.

Preferred host cells for expressing the CD antigen variants of this invention are mammalian cell lines, examples including monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651), human embryonic kidney line (293, Graham, F.L. et al. *J. Gen. Virol.* 36: 59 (1977) and 293s cells [293 subclones selected for better suspension growth]), baby hamster kidney cells (BHK, ATCC CCL 10), chinese hamster ovary-cells-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. (USA)* 77: 4216, [1980]), mouse sarcoma cells (TMA, Mather, J.P. *Biol. Reprod.* 23: 243-251 [1980]), monkey kidney cells (CV1 ATCC CCL 70), african green monkey kidney cells (VERO-76, ATCC CRL-1587), human cervical carcinoma cells (HELA, ATCC CCL 2), canine kidney cells (MDCK, ATCC CCL 34), buffalo rat liver cells (BRL 3A ATCC CRL 1442), human lung cells (W138, ATCC CCL 75), human liver cells (Hep G2, HB 8065), mouse mammary tumor (MMT 060562, ATCC CCL51 cells), and TRI cells (Mather, J.P. et al., *Annals N.Y. Acad. Sci.* 383: 44-68 [1982]).

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. One suitable for transformation of the host cells is the method of Graham, F. and van der Eb, A., *Virology* 52: 456-457 (1973). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell walls are used as hosts, the preferred method of transformation is calcium treatment using calcium chloride as described by Cohen, F.N. et al., *Proc. Natl. Acad. Sci. (USA)*, 69: 2110 (1972).

Construction of suitable vectors containing the desired coding and control sequences employ standard and manipulative ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required. Suitable procedures are well known for the construction described herein. See for example, (Maniatis, T. et al., *Molecular Cloning*, 133-134 Cold Spring Harbor, [1982], "Current Protocols in Molecular Biology" edited by Ausubel et al., [1987], pub. by Greene Publishing Associates & Wiley-Interscience).

Correct plasmid sequences are confirmed by transforming *E. coli* K12 strain 294 (ATCC 31446) with ligation mixtures, successful transformants selected by ampicillin or tetracycline resistance where appropriate, plasmids from the transformants prepared, and then analyzed by restriction enzyme digestion and or sequences by the method of Messing et al., *Nucleic Acids Res.* 9: 309 (1981) or by the method of Maxam et al., *Methods in Enzymology* 65: 499 (1980).

Host cells are transformed with the expression vectors of this invention. Thereafter they are cultured in appropriate culture media, e.g. containing substances for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The secreted adhesion variants are recovered and purified from the culture supernatants of recombinant hosts. Typically, the supernatants are concentrated by ultrafiltration, contacted with a ligand affinity or immunoaffinity matrix so as to adsorb the adhesion variant, and eluted from the matrix. Optionally, the adhesion is purified by ion exchange chromatography.

Surprisingly purification of soluble CD4 adhesion from culture medium was unexpectedly difficult. Notwithstanding that the hydrophobic transmembrane region of the antigen had been deleted, the antigen exhibited a strong tendency to form aggregates that could be readily removed from suspension by centrifugation at 1000 x g and which avidly coat surfaces such as ultrafiltration membranes. This appears to result from the reduction in concentration of albumin or other serum protein (ordinarily present in the crude preparation) to a particular level, below which the truncated antigen no longer remains soluble. This phenomenon appears to be aggravated by exposure of the CD4 adhesion to low pH (< about pH 4). As a result, separation procedures (particularly those that employ acid elution, such as immunoaffinity) should be modified so that the eluate is maintained at, or immediately returned to, about neutrality. Further, a surfactant, e.g. a detergent such as Tween 80, should be included with the antigen during the separation procedure. The final purified product will be stabilized with a predetermined protein such as albumin, and or a detergent.

The purified adhesion is formulated into conventional pharmacologically acceptable excipients.

It is administered to patients having HIV infection at a dosage capable of maintaining a concentration of greater than about 100 ng of soluble CD4 adhesion/ml plasma. For CD4 adhesion variants having different molecular weights, about 2 picomoles of soluble receptor per ml of plasma will be initially evaluated clinically in order to establish a stoichiometric equivalence with native (membrane bound) and soluble receptor. The ordinary dosage of soluble CD4 is 100 µg/kg of patient weight/day.

The therapeutic CD4 variants are employed with other therapies and agents for the treatment of AIDS including AZT, neutralizing antibodies and immunocytotoxins, gp120 fragments and vaccines.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes typically 1 µg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 µl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally (Lawn, R. et al., *Nucleic Acids Res.* 9: 6103-6114 [1981], and Goecdel, D. et al., *Nucleic Acids Res.* 8: 4057 [1980]).

"Dephosphorylation" refers to the removal of the terminal 5' phosphates by treatment with bacterial alkaline phosphatase (BAP). This procedure prevents the two restriction cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Procedures and reagents for dephosphorylation and other recombinant manipulations are conventional. Reactions using BAP are carried out in 50 mM Tris at 68°C to suppress the activity of any exonucleases which may be present in the enzyme preparations. Reactions were run for 1 hour. Following the reaction the DNA fragment is gel purified.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T. et al. *Id.* at 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 µg of approximately equimolar amounts of the DNA fragments to be ligated.

"Filling" or "blunting" refers to the procedures by which the single stranded end in the cohesive terminus of a restriction enzyme-cleaved nucleic acid is converted to a double strand. This eliminates the cohesive terminus and forms a blunt end. This process is a versatile tool for converting a restriction cut end that may be cohesive with the ends created by only one or a few other restriction enzymes into a terminus compatible with any blunt-cutting restriction endonuclease or other filled cohesive terminus. Typically, blunting is accomplished by incubating 2-15 $\mu$ g of the target DNA in 10mM MgCl<sub>2</sub>, 1mM dithiothreitol, 50mM NaCl, 10mM Tris (pH 7.5) buffer at about 37°C in the presence of 8 units of the Klenow fragment of DNA polymerase I and 250  $\mu$ M of each of the four deoxynucleoside triphosphates. The incubation generally is terminated after 30 min phenol and chloroform extraction and ethanol precipitation.

The following examples merely illustrate the best mode now contemplated for practicing the invention, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

### Example 1

#### Construction of Vectors for the Expression of Native CD4 and Secreted Derivatives

##### Section 1

The plasmid used for recombinant synthesis of human CD4 was pSveCD4DHFR. The plasmid was constructed as follows.

$\lambda$ CD4P1 containing most of the coding sequence of human CD4 (obtained from a human placental cDNA library using oligonucleotide probes based on the published sequence (Maddon et al. 1985)) was digested with EcoRI to produce the cDNA insert. This fragment was recovered by polyacrylamide gel electrophoresis (fragment 1).

pUC18 was digested with EcoRI and the single fragment recovered by polyacrylamide gel electrophoresis (fragment 2). Fragment 1 was ligated to fragment 2 and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct DNA fragments. This plasmid is referred to as pUCCD4.

pSveE DHFR (Muesing et al., Cell 48:691-701 (1987)) was digested with KpnI and BamHI and blunted with E. coli DNA polymerase I (Klenow fragment) and the four dNTPs. Fragment 3 containing the pML-Amp' region, SV40 early promoter, the HIV LTR, and the mouse DHFR gene was recovered by gel electrophoresis, ligated and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the BamHI restriction site and the absence of the KpnI restriction site. This plasmid is referred to as pSve $\Delta$ BKDHFR and allows EcoRI-BamHI fragments to be inserted after the SV40 early promoter and transcribed under its control, following transfection into an appropriate cell line.

Synthetic oligonucleotides (adaptors 1-8, below) were made to extend from 76 bp 5' of the initiation codon of CD4 translation to the RsaI restriction site at 121 bp 3' of the initiator, with the sequence AATT at the 5' end of the sense strand to generate an end which could ligate to an EcoRI restriction fragment. These oligonucleotides were ligated and the 204 bp fragment containing the entire sequence recovered by gel electrophoresis (fragment 4).

CD4 adaptor 1: AATTCAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCCT  
 CD4 adaptor 2:  $\alpha$ ACTGCTCAGCCCCCTTCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGTC  
 CD4 adaptor 3: pCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCAGC  
 adaptor  
 pAGCCACTCAGGGAAACAAAGTGGTGCTGGGCAAAAAGGGGATACAGTGGAAGTACCTGT  
 CD4 adaptor 5: pACAGGTCAAGTCCACTGTATCCCCTTTTGGCCAGCACCAGTGTGTTCC  
 CD4 adaptor 6: pCTGAGTGGCTGCTGGGAGGAGCGCCAGTTGCAGCACCAGAAGCAAGT  
 CD4 adaptor 7: pGCCTAAAAGGGACTCCCCGGTTCATTGTGGCCTTGCCGAGGGAGGAAGGG

CD4 adaptor 8 GCTGAGCAGTAGGGACCTGAGCCACAGAAATGGCAGGGCTCTGGGCTTG

pUCCD4 was digested with RsaI and SstI and the 401 bp fragment containing part of the CD4 coding sequence recovered by gel electrophoresis (fragment 5). pUC18 was digested with EcoRI and SstI and the fragment comprising the bulk of the plasmid recovered by gel electrophoresis (fragment 6). Fragments 4 and 5 were ligated to fragment 6 and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of the inserted synthetic DNA was checked by excising the 605 bp EcoRI-SstI fragments from several transformants and ligating them to M13mp19 which had been digested with the same enzymes. After transformation into E. coli strain JM101, single-stranded DNA was prepared and sequenced. One plasmid which contained the correct sequence was selected, and is referred to as pCD4int.

pCD4int was digested with EcoRI and SstI and fragment 7 containing the 5' end of the CD4 coding region was recovered by gel electrophoresis. pUCCD4 was digested with SstI and BamHI and the 1139 bp fragment containing the remainder of the CD4 coding region (fragment 8) recovered by gel electrophoresis.

pSV $\Delta$ BKDHFR was digested with EcoRI and BamHI and fragment 9 comprising the bulk of the plasmid was isolated. Fragments 7, 8 and 9 were ligated and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and the resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSV $\Delta$ CD4DHFR, and was used to direct synthesis of recombinant intact CD4.

## Section 2

A plasmid was constructed to direct the synthesis of a CD4 derivative lacking the putative transmembrane domain and most of the putative cytoplasmic domain (Maddon et al.) This was done with the intention of creating a secreted form of CD4, based on the assumption that these domains anchor the CD4 glycoprotein to the cell membrane, and that their deletion would result in the secretion of the product. This plasmid is referred to as pSV $\Delta$ CD4 $\Delta$ N1aDHFR and was constructed as follows:

pUCCD4 was digested with SstI and TaqI and the 531 bp fragment (fragment 10) recovered. pUCCD4 was digested with NlaIII and TaqI and the 112 bp fragment (fragment 11) recovered. pUCCD4 was digested with BamHI and NlaIII and the 301 bp fragment (fragment 12) recovered. pCD4int was digested with SstI and BamHI and fragment 13 comprising the bulk of the plasmid recovered. Fragments 10, 11, and 12 were ligated together with fragment 13 and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. Plasmid DNA from several transformants was sequenced to ensure that the 195 bp NlaIII fragment had been deleted and that the proper reading frame was restored. The resulting plasmid is referred to as pCD4 $\Delta$ N1a.

pCD4 $\Delta$ N1a was digested with EcoRI and BamHI and the 1541 bp fragment containing the sequence of a CD4 derivative lacking the transmembrane and cytoplasmic domains recovered (fragment 14) and ligated to fragment 9 and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSV $\Delta$ CD4 $\Delta$ N1aDHFR.

Both pSV $\Delta$ CD4DHFR and pSV $\Delta$ CD4 $\Delta$ N1aDHFR were transfected into CHO cells by the same method used to establish cell lines stably expressing HIV-1 polypeptides (Muesing, Smith and Capon, Cell 48:6910701 [1987]). These cells were assayed for production by radioimmuno-precipitation as described below. While no product was detected in initial experiments, subsequent experiments showed that the above described coding segment could indeed direct the synthesis of a soluble CD4 adhesion variant both in CHO and 293 cells.

## Section 3

A different expression system was initially used for the synthesis and expression of a CD4 variant lacking completely the cytoplasmic and transmembrane domains. The system uses the cytomegalovirus

promoter and can be used in cultured cells of human origin. The first plasmid constructed for use in this system contained the entire coding region for CD4 and was intended to function as a control in the following studies. It is referred to as pRKCD4, and was constructed as follows:

pSVeCD4DHFR was digested with EcoRI and BamHI and fragment 15 containing the entire CD4 coding region was isolated. pRK5 (European Application No. 88308386.7) was digested with EcoRI and BamHI and fragment 16 comprising the bulk of the plasmid recovered by gel electrophoresis, ligated to fragment 15, and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4.

#### Section 4

The next plasmid constructed was designed to direct the expression of the above-mentioned (Section 3) secreted derivative of CD4. The coding region of CD4 was fused after amino acid residue 368 of mature CD4 to a sequence from pBR322 which codes for 9 more residues before a translation termination codon. This removes the putative CD4 transmembrane and cytoplasmic domains, which are presumed to anchor CD4 to the cell surface. The plasmid is referred to as pRKCD4T, and was constructed as follows:

pSVeCD4DHFR was digested with HpaII, blunted with Klenow fragment and the four dNTPs, and digested with BstEII. The 382 bp fragment (fragment 17) containing part of the CD4 coding sequence was recovered by gel electrophoresis. pSVeCD4DHFR was digested with EcoRI and BstEII and the 874 bp fragment (fragment 18) recovered. pBR322 was digested with HindIII, blunted with Klenow fragment and the four dNTPs, and digested with EcoRI. Fragment 19 comprising the bulk of the plasmid was isolated and ligated to fragments 17 and 18 and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pCD4Tint.

pRK5 was digested with EcoRI and SmaI and fragment 20 comprising the bulk of the plasmid isolated. pCD4Tint was digested with EcoRI and EcoRV and the 1410 bp fragment containing the CD4 coding sequence to the HpaII site at 1176 bp 3' of the initiating codon and the 154 bp HindIII-EcoRV fragment of pBR322 was recovered (fragment 21). Fragments 20 and 21 were ligated and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4T.

#### Section 5a

In order to create a secreted form of CD4 which could be purified with an antibody directed to herpes virus type I glycoprotein D, a plasmid was constructed to express a derivative of CD4T in which the region coding for the mature, processed CD4T polypeptide was fused to a sequence coding for the signal peptide and the first 27 residues of the mature type I Herpes Simplex Virus gD glycoprotein. This plasmid is referred to as pRKGD4T, and was constructed as follows:

pgDTrunc.DHFR was digested with EcoRI and PvuII and the fragment containing the coding region for the signal peptide and first 27 residues of the mature HSV I gD glycoprotein was isolated (fragment 22). pRKCD4T was digested with EcoRI and BstEII and fragment 23 containing the 3' end of the CD4 coding sequence and the pRK5 region was isolated.

Synthetic oligonucleotides GD (adaptors 1-2, below) containing the coding sequence of CD4 from the codon for the amino terminal residue of mature CD4 to the Rsa site at 121 bp 3' of translation initiation, and containing the sequence CTGCTCGAG at the 5' end of the sense strand were prepared (fragment 24). pRKCD4 was digested with RsaI and BstEII and the 665 bp fragment containing part of the coding region for CD4 was recovered (fragment 25) and ligated to fragment 24. After digestion with BstEII to ensure that only monomeric fragment was present, the 724 bp fragment containing both sequences was recovered by gel electrophoresis (fragment 26).

Fragments 22, 23 and 26 were ligated and the ligation mixture transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of several transformants was checked to ensure that the synthetic insert was



correct and that reading frame was preserved. This plasmid is referred to as pRKGD4T.

These pRK5 derived plasmids preferably were transfected into 293S cells for stable expression according to Muesing, et al. Cell 48:691 (1987) with the exception that in addition to the plasmid of interest a plasmid expressing the neomycin resistance gene pRSV neo (Gorman et al. Science 221:553-555 (1985)) was cotransfected. 293 cells also are used satisfactorily as host cells. 2 days after transfection, the cells were passaged into standard medium (1:1 F12 DME supplemented with L-glutamine, penicillin-streptomycin and 10% FBS) with 0.5 mg/ml G-418 (Genticin sulfate; Gibco) for selection of stable cell lines, rather than in media containing methotrexate as shown by Muesing et al. Cells were assayed for production of CD4 or CD4 analogs by radioimmunoprecipitation. Binding studies (section 5c) used conditioned supernatants from these cells in the 1:1 F12 DME medium. Materials used in infectivity assays (section 5b) were obtained as described in section 8 below.

```

gDCD4                                adaptor                                1:
CTGCTCGAGCAGGGAAACAAAGTGGTGGCTGGGCAAAAAGGGGATACAGTGGAAGTAC
gDCD4                                adaptor                                2:
pACAGGTCAGTTCCTGTATCCCCCTTTTGGCCAGCACCCTTTGTTCCCTGCTCGA

```

### Section 5b

The following constitutes a study of the neutralization of HIV-1 infectivity by soluble CD4 analogs. A modification of the neutralization procedure of Robert-Guroff et al., Nature 316:72 (1985) was followed. Equal volumes of inhibitor supernatant and virus (60 microliters) were incubated at 4 degrees C for 1 hour, then the same volume of H9 (Gallo et al., Science 224:500, 1984) at  $5 \times 10^5$  ml was added and incubation continued for 1 hour at 37 degrees C. Following absorption,  $2.5 \times 10^5$  cells in 150 microliters were transferred to 2 ml of incubation media. After 4 days at 37 degrees C, the cultures were split 1:2 with fresh media and incubated for an additional 3 days. Cultures were harvested, reverse transcriptase activity was measured (Groopman et al., AIDS Research and Human Retroviruses 3:71, 1987), and immunofluorescence reactivity with HIV-1 positive serum was determined as described (Poesz et al., Proc. Acad. Nat. Sci. USA 77:7415, 1980). Inhibitor supernatants were obtained from confluent plate cultures of 293s CDT4, 293s gDCD4T cells or untransfected 293s cells by replacing the growth medium incubation media and harvesting the supernatants 24 hours later. Inhibitor supernatant replaced part or all of the incubation media during the first three days of culture as indicated in the second column of Table 3. Challenge dose of virus was 100 TCID<sub>50</sub> (Groopman et al., supra) of HIV-1 strain HTLV-IIIb grown in H9 cells assayed in the same system. Incubation media consisted of RPMI 1640 media containing 2mM L-glutamine, 100 units/ml penicillin, 100 micrograms/ml streptomycin, 2 micrograms/ml polybrene and 20% fetal calf serum (M.A. Bioproducts).

Table 3

Inhibitor supernatant	Dilution of Inhibitor supernatant	Indirect immunofluorescence (% positive cells)		Reverse transcriptase (cpm/ml $\times 10^5$ )	
mock-transfected	undil., 1:4	65.3	65.5	21.8	23.9
mock-transfected	undil., 1:4	61.2	61.1	18.5	28.1
CD4T	undil., 1:4	0.4	18.0	0.11	5.94
CD4T	undil., 1:4	0.8	16.1	0.15	3.72
gDCD4T	undil., 1:4	0.4	26.8	0.14	9.92
gDCD4T	undil., 1:4	1.4	36.1	0.23	11.3

Both forms of soluble CD4 virtually abolished the growth of HIV-1, when incubated with virus-infected cells without prior dilution (Table 2). At a dilution of 1:4 the soluble CD4 preparations were only partially effective in inhibiting virus growth, however the level of fluorescent-positive cells and reverse transcriptase was still significantly lower than cultures receiving mock-transfected cell supernatants (Table 2). Since there was no significant difference in virus growth between diluted and undiluted control supernatants, nor did any of the supernatants affect the growth of uninfected H9 cells (data not shown), soluble CD4 proteins present in these supernatants were concluded to be responsible for the neutralization of HIV-1 infection of H9 cells.



### Section 5c

To determine the affinity constant for interactions between gp120 and CD4 or CD4 variants, saturation binding analysis was carried out with soluble CD4 (supra) and detergent solubilized intact CD4 (Lasky et al. Cell 50:975 (1987)) employing radiolabeled gp120 labeled with lactoperoxidase. Binding reactions consisted of  $^{125}$ I-gp120 (3 ng to 670 ng, 2.9 nCi/ng) incubated for 1 hour at 0 degrees C with cell lysates containing intact CD4 (Lasky et al., op cit.) or cell supernatants containing unlabeled CD4T or gDCD4T prepared as described in section 5a. Reactions (0.2ml) had a final composition of 0.5X McDougal Lysis Buffer (McDLB) (1 x McDLB contains 0.5% Nonidet NP-40, 0.2% Na deoxycholate, 0.12 M NaCl, 0.02 M Tris-HCl, pH 8.0) and were performed in duplicate, both in the presence or absence of 50 micrograms of unlabeled purified gp120 (74 fold or greater excess). Following incubation, bound gp120 was quantitated by immunoprecipitation and counted in a gamma counter. For immunoprecipitation, binding reaction solutions were preabsorbed with 5 microliters of normal rabbit serum for one hour at 0 C, and cleared with 40 microliters of Pansorbin (10% w/v, Calbiochem) for 30 minutes at 0 degrees C. Samples were then incubated overnight at 0 degrees C with 2 microliters of normal serum or 5 microliters (0.25 microgram) of OKT4 monoclonal antibody (Ortho) followed by collection of immune complexes with 10 microliters of Pansorbin. Precipitates were washed twice in 1X McDLB and once in water, then eluted by eluting at 100 degrees C for 2 minutes in sample buffer (0.12 M Tris-HCl pH 6.8, 4% SDS, 0.7 M mercaptoethanol, 20% glycerol, and 0.1% bromophenol blue). CD4 molecules were bound saturably by gp120, and yielded a simple mass action binding curve. Supernatants from mock-transfected cells gave a level of specifically bound gp120 less than 1% that found for supernatants containing soluble CD4. Scatchard analysis revealed a single class of binding sites on each molecule, with apparent dissociation constants (Kd) of  $1.3 \times 10^{-7}$  M,  $0.83 \times 10^{-7}$  M and  $0.72 \times 10^{-7}$  M for intact CD4, CD4T and gDCD4T, respectively. The values obtained for CD4-gp120 binding in solution are comparable to the affinity previously measured for gp120 binding to CD4 on whole cells (Kd =  $4.0 \times 10^{-7}$  M Lasky, Cell, supra).

### Section 6

In order to produce secreted derivatives of CD4 which are free of extraneous amino acid residues, two plasmids were constructed for expression in 293 cells. The plasmids contain CD4 genes which have been truncated without the addition of extra residues, and are referred to as pRKCD4ΔN1a and pRKCD4TP, and were constructed as follows:

Fragment 14 containing the CD4 gene with the 195 bp N1aIII restriction fragment deleted was ligated to fragment 16, which is pRK5 digested with EcoRI and BamHI. The ligation mixture was transformed into E. coli strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4ΔN1a.

Synthetic DNA was made to attach to the HpaII site at 1176bp and which when so attached would terminate translation after amino acid residue 370 of mature CD4 (fragment 27). The other end of this fragment was designed to ligate to BamHI restriction fragments. pUCCD4 was digested with BstEII and HpaII and the 382bp fragment containing part of the CD4 gene was recovered (fragment 28). Fragments 27 and 28 were ligated and then digested with BstEII to reduce dimerized fragments to monomers, and the resulting 401bp fragment was recovered (fragment 29).

pRKCD4 was digested with BstII and BamHI and the fragment comprising the bulk of the plasmid (fragment 30) was isolated and ligated to fragment 29. The ligation mixture was transformed into E. coli strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4TP. Both plasmids are transfected into 293 cells to generate stable variant CD4-expressing cell lines as described above.

### Section 7

Two plasmids were constructed to direct the expression of secreted CD4 lacking extraneous amino acid residues in CHO cells. These are referred to as pSveCD4ΔN1aSVDHFR and pSveCD4TPSVDHFR, and were constructed as follows:

pE348HBV.E400D22 was digested with PvuI and EcoRI and the fragment containing the SV40 early

promoter and part of the  $\beta$ -lactamase gene was recovered (fragment 31). pE348HBV E400D22 was digested with PvuI and BamHI and the large fragment containing the balance of the  $\beta$ -lactamase gene as well as the SV40 early promoter and the DHFR gene was isolated (fragment 32).

Fragments 31 and 32 were ligated together with fragment 14 and transformed into E. coli strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVECD4 $\Delta$ N1aSVDHFR. This plasmid contains the same DNA fragment encoding the soluble CD4 molecule found in the above-mentioned plasmid pSVECD4 $\Delta$ N1aDHFR (Section 2).

pRKCD4TP was digested with EcoRI and BamHI and the fragment containing the truncated CD4 coding region was isolated and ligated to fragments 31 and 32. The ligation mixture was transformed into E. coli strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVECD4TPSVDHFR. Both of these plasmids are transfected into CHO cells and amplified transfectants selected by methotrexate using conventional procedures.

### Example 2

Fusions of the V region of the CD4 gene which is homologous to the variable region of immunoglobulin genes (ref Maddon et al. 1985), to the constant (C) region of human immunoglobulin  $\alpha$  and  $\gamma$ 2 chains are constructed as follows:

Synthetic DNA is made to code for the C region of human  $\alpha$  chain (residues 109-214) based on the sequence published by Morin et al., Proc. Natl Acad Sci 82:7025-7029, with the addition at the 5' end of the coding strand of the sequence GGGG, which allows this fragment to be ligated to the BspMI site at the end of the putative V-like region of CD4. At the 3' end of the coding region, a translational stop codon is added as well as a sequence which allows this end to be ligated to BamHI restriction fragments. The synthetic DNA is made in 8 fragments, 4 for each strand, 70-90 bases long. These are then allowed to anneal and ligated prior to isolation on a polyacrylamide gel (fragment 33).

pRKCD4 is digested with EcoRI and BspMI and the 478bp fragment containing the region coding for the putative V-like domain of CD4 is recovered (fragment 34). Fragments 33 and 34 are ligated together with fragment 16 (from the expression vector pRK5). The ligation mixture is transformed into E. coli strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA is prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4C $\alpha$ .

A plasmid encoding a fusion of the CD4 V-like domain to the human immunoglobulin  $\gamma$ 2 region is constructed in a similar fashion, and is referred to as pRKCD4C $\gamma$ 2. Both of these plasmids are transfected into 293 cells, myeloma cells or other competent cells in order to obtain cell lines expressing variant CD4 molecules as described above.

### Expression in CHO Cells

Plasmids were constructed to direct the expression of the immunoadhesons described above in CHO cells. These are referred to as pSVECD4 $\alpha$  $\gamma$ 1SVDHFR, pSVECD4 $\gamma$ 1SVDHFR, pSVECD4 $\alpha$  $\gamma$ 1SVDHFR, pSVECD4 $\alpha$  $\gamma$ 1SVDHFR, pSVECD4 $\alpha$  $\gamma$ 1SVDHFR and pSVECD4 $\gamma$ 1SVDHFR.

Fragment 31 was prepared as described above. Fragment 32a was prepared by digesting plasmid pE348HBV E400 D22 with BamHI, blunting with Klenow fragment and the four dNTPs, then digesting with PvuI. Plasmids pRKCD4 $\alpha$  $\gamma$ 1, pRKCD4 $\gamma$ 1, pRKCD4 $\alpha$  $\gamma$ 1, pRKCD4 $\gamma$ 1, pRKCD4 $\alpha$  $\gamma$ 1, and pRKCD4 $\gamma$ 1 were separately digested with HindIII, blunted with Klenow fragment and the four dNTPs, then digested with EcoRI. The resulting DNA fragments were ligated together with fragments 31 and 32a and transformed into E. coli strain 294. Colonies were selected and checked for the presence of the correct plasmid as above, then transfected into CHO cells and amplified by methotrexate selection using conventional procedures.

### Example 3

The gDCD4T secreted by the method of Example 1 was purified from cell culture fluid containing either 10% FBS (fetal bovine serum) or no added FBS. The conditioned cell culture fluid was first concentrated by ultrafiltration then purified by immunoaffinity chromatography. The immunoaffinity column was produced by coupling murine monoclonal antibody 5B6 (whose epitope is on the HSV-1 gD portion of the gDCD4T molecule) to glyceryl coated controlled pore glass by the method of Roy et al. 1984. The concentrated cell culture fluid is applied directly to the column and the contaminating proteins are washed away with neutral pH buffer. The column is then washed with neutral buffer containing tetramethylammonium chloride followed by neutral buffer containing Tween 80. The bound gDCD4T is eluted from the column with buffer at pH3 containing Tween 80 (0.1% w/v) and is neutralized immediately as it is eluted. The eluted neutralized gDCD4T is then concentrated by ultrafiltration and dialyzed/diafiltered to exchange the buffer for a physiological salt solution containing Tween 80 at approximately 0.1% w/v.

If the detergent is not present the gDCD4T forms aggregates as evidenced by the ability of centrifugation at approximately 10,000 Xg for 2 minutes to remove the gDCD4T from the solution. Incubation of gDCD4T at 4°C in 0.1M sodium acetate, 0.5M NaCl and 0.25M Tris at pH 7 together with BSA, Tween 80 or glycerol as candidate stabilizers showed that, in the absence of a stabilizer the gDCD4T gradually aggregated over the space of 12 days to the point where only about 60-70% of the protein was soluble. However, use of 0.1% w/v Tween 80 or (0.5 mg/ml BSA ensured that about 100% or 80%, respectively, of the gDCD4T remained soluble over this period. Surprisingly glycerol was ineffective as a stabilizer and produced results inferior even to the control - at 8 days about 80% of the gDCD4T was aggregated when stored in the presence of glycerol.

#### Example 4

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin  $\gamma 1$ . These plasmids are referred to as pRKCD4<sub>2,1</sub>, pRKCD4<sub>4,1</sub>, pRKCD4<sub>2,1</sub>, pRKCD4<sub>2,1</sub>, pRKCD4<sub>1,1</sub>, and pRKCD4<sub>1,1</sub>.

Plasmid pRKCD4<sub>1,1</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin  $\gamma 1$ , starting at the codon for serine residue 114 of mature human immunoglobulin  $\gamma 1$  (Kabat et al.).

Plasmid pRKCD4<sub>4,1</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin  $\gamma 1$ , starting at the codon for serine residue 114 of mature human immunoglobulin  $\gamma 1$  (Kabat et al.).

Plasmid pRKCD4<sub>2,1</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for glutamine residue 180 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin  $\gamma 1$ , starting at the codon for serine residue 114 of mature human immunoglobulin  $\gamma 1$  (Kabat et al.).

Plasmid pRKCD4<sub>2,1</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 177 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin  $\gamma 1$ , starting at the codon for serine residue 114 of mature human immunoglobulin  $\gamma 1$  (Kabat et al.).

Plasmid pRKCD4<sub>1,1</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for aspartic acid residue 105 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin  $\gamma 1$ , starting at the codon for serine residue 114 of mature human immunoglobulin  $\gamma 1$  (Kabat et al.).

Plasmid pRKCD4<sub>1,1</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 100 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin  $\gamma 1$ , starting at the codon for serine residue 114 of mature human immunoglobulin  $\gamma 1$  (Kabat et al.).

Construction of these plasmids required the prior construction of plasmid pRKCD4TP $\gamma 1$ . It was constructed as follows:

A cDNA clone coding for human immunoglobulin  $\gamma 1$  was obtained from a human spleen cDNA library

(Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Elison *et al.*, "Nucleic Acids Res." 10 4071-4079 [1982]), and an EcoRI-EagI fragment (the EcoRI site was contributed by a linker) containing part of the variable and all of the constant region was obtained. This fragment was blunted with Klenow fragment, and recovered by gel electrophoresis (Fragment a1).

Plasmid pRKCD4TP was digested with XbaI and treated with Klenow Enzyme, and Fragment a2, containing the linearized plasmid was recovered by gel electrophoresis, and ligated with fragment a1. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from the transformants and checked by restriction analysis for the presence of the correct fragment in the correct orientation (i.e. the immunoglobulin coding region in the same orientation as the CD4 coding region, and at the 3' end of the CD4 coding region). This plasmid is referred to as pRKCD4TP  $\gamma$ 1.

Synthetic oligonucleotides were made as primers for deletional mutagenesis reactions to fuse the appropriate coding sequences of IgG1 and CD4 as described above. These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH<sub>2</sub>-terminal 8 residues of the desired immunoglobulin moiety). Plasmid pRKCD4TP  $\gamma$ 1 was transformed into *E. coli* strain SR101 and the transformed cultures plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13KO7 bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP  $\gamma$ 1. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13KO7 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmids were transfected into 293 cells using standard procedures and assayed for expression and production as described above.

	Expressed	Secreted
pRKCD4 $\gamma$ 1	-	-
pRKCD4 $\gamma$ 1	+	-
pRKCD4 $\gamma$ 1	+	+
pRKCD4 $\gamma$ 1	+	+
pRKCD4 $\gamma$ 1	+	+
pRKCD4 $\gamma$ 1	+	+

Plasmids also were constructed to direct the expression of fusion proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the truncated portion of the constant region of human immunoglobulin  $\gamma$ 1, comprising only the hinge region and constant domains CH<sub>2</sub> and CH<sub>3</sub>.

Synthetic oligonucleotides were made as primers for mutagenesis reactions to create the immunoglobulin sequence from Ser114 to Cys215 inclusive (Kabat *et al.*). These were synthesized as 48-mers comprising 242 $\frac{1}{2}$  nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH<sub>2</sub>-terminal 8 residues of the desired immunoglobulin moiety). Plasmids pRKCD4 $\gamma$ 1, pRKCD4 $\gamma$ 1 and pRKCD4 $\gamma$ 1 were separately transformed into *E. coli* strain SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (Grunstein-Hogness) for the presence of the appropriate fusion site using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13KO7 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmid derived from plasmid pRKCD4 $\gamma$ 1 is referred to as pRKCD4 $\gamma$ 1, that derived from plasmid pRKCD4 $\gamma$ 1 is referred to as pRKCD4 $\gamma$ 1, and that derived from plasmid pRKCD4 $\gamma$ 1 is referred to as

pRKCD4<sub>ser</sub>.

pRKCD4<sub>ser</sub>, pRKCD4<sub>lys</sub> and pRKCD4<sub>trc</sub> are cultured in the same fashion as described above and CH1-deleted CD4 immunoadhesons recovered as described elsewhere herein

### Light Chain Fusions

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin  $\kappa$ . These plasmids are referred to as pRKCD4<sub>ser</sub> and pRKCD4<sub>lys</sub>.

Plasmid pRKCD4<sub>ser</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin  $\kappa$ , starting at the codon for threonine residue 109 of the mature human immunoglobulin  $\kappa$ . (Kabat *et al.*)

Plasmid pRKCD4<sub>lys</sub> contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin  $\kappa$ , starting at the codon for threonine residue 109 of the mature human immunoglobulin  $\kappa$ . (Kabat *et al.*)

These plasmids were constructed in a manner analogous to plasmid pRKCD4<sub>trc</sub> described above, with the following exception.

The human immunoglobulin  $\kappa$  coding sequence (Fig. 5) was obtained from a human spleen cDNA library (Clontech Laboratories Inc.) using oligonucleotides based on the published sequence (Heter, P.A. *et al.*, Cell 22:197-207 [1980]) and an *EcoRI*-*BspMI* fragment containing part of the variable region and the entire constant region was obtained. This fragment was blunted with Klenow fragment and the four dNTPs. This fragment was used instead of fragment  $\alpha$ , and was used to construct plasmid pRKCD4TP h.

### Example 5

#### Culture, Purification and formulation of CD4 variants

Plasmids encoding CD4T, prolyl terminal (CD4TP), or CD4T immunoadhesons were calcium phosphate transfected into CHO-DP7 (a proinsulin-transformed autocrine host cell derived from CHO; U.S.N. 97 472) and the transformants grown in selective medium (1:1 HAM F12 DMEM GHT<sup>-</sup> containing 1 - 10% dialyzed or dialyzed bovine serum). Other suitable host cells are CHO cells or 293s human embryonic kidney cells. The transformants were subcloned into the same medium but containing 500 nm methotrexate. A subclone capable of secreting CD4T, CD4tp 500 b, was selected. CD4tp 500 b is cultured in a DMEM HAM F12 medium at about 37°C until CD4T accumulates in the culture, after which the medium is separated from the cells and insoluble matter by centrifuging.

Culture fluid from CD4TP transformants was concentrated and dialyzed to lower the ionic strength. The concentrate was passed through a large volume of Q-Sepharose anion exchange resin (previously equilibrated with 25 mM NaCl, pH 8.5) in order to adsorb contaminants from the culture fluid. The isoelectric point of CD4TP is about 9.5, thus making it possible to discriminate between truncated forms of CD4 and most contaminants by alternate adsorption, respectively, on a cation exchange resin such as carboxymethyl or sulfonyle Sepharose, and an anion exchange resin such as a quaternary ammonium Sepharose. In addition, since highly electropositive domains are present in the extracellular segment of CD4 and CD4-containing variant is purified in the same fashion as CD4TP. The unadsorbed culture fluid from the anion exchange resin step was then passed through a cation exchange resin (previously equilibrated with 25 mM NaCl at pH 8.5) whereby CD4TP was adsorbed to the resin. The CD4TP was eluted with a NaCl gradient at pH 8.5, this CD variant eluting at about 0.2 mM NaCl. Ammonium sulfate was added to the eluate to the concentration of 1.7M and the solution passed through a column of hydrophobic interaction chromatography resin (phenyl or butyl Sepharose). The CD4TP was eluted from the hydrophobic interaction column with a gradient of ammonium sulfate, the CD4TP emerging at about 0.7M ammonium sulfate. The eluate was concentrated and buffer exchanged on a G-25 column using phosphate buffered saline containing 0.2% (w/v) Tween 20 or Tween 80. The CD4TP was soluble and stable in this solution, which

was sterile filtered and filled into vials as an aqueous formulation. Other polymeric nonionic surfactants are suitably used with the CD4 formulations, including Pluronic block copolymers or polyethylene glycol.

It is also possible to employ immunoaffinity purification of soluble CD4 wherein the CD4 is adsorbed onto an immobilized antibody against CD4. This method suffers from the disadvantage that elution of the CD4T under acidic conditions leads to protein aggregation that is only thoroughly ameliorated at relatively higher levels of surfactant. The foregoing procedure permits the use of much lower quantities of surfactant, about from 0.01 to 0.10 % (w/v) surfactant.

The procedure followed for the purification of CD4 fusions with immunoglobulin heavy chain was to concentrate recombinant supernatants by ultrafiltration and thereafter adsorb the fusion onto resin-immobilized Staphylococcal protein A. The fusion was eluted with 0.1M citrate buffer pH 3 with no salt or detergent. This preparation is buffered into Tris buffer at pH 7.5. The immunoglobulin fusions with CD4 V1-V4 optionally are further purified by the procedure described above for unfused CD4 variants. CD4 immunoglobulin fusions with CD4 V1-V2 also may be purified by the procedure above, except that it is not expected that the isoelectric point of this class of molecules will be as alkaline as that of species containing all four V regions of CD4.

#### Example 6

The characteristics of several adhesion variants were determined. As shown in table 4 the immunoadhesons CD4<sub>471</sub> and CD4<sub>771</sub> show improved plasma half-life in rabbits, coupled with high-affinity gp120 binding and an affinity for Fc<sub>γ</sub> receptor (determined with U937 cells) that is comparable to that of bulk human IgG.

Table 4

	gp120 KD (nM) <sup>a</sup>	Fc <sub>γ</sub> R KD (nM) <sup>b</sup>	Plasma Half-Life <sup>c</sup> in Rabbits (Hrs.)
CD4T <sup>d</sup>	2.3 ± 0.4	-	0.25
CD4 <sub>471</sub>	1.2 ± 0.1	2.83 ± 0.25	6.4
CD4 <sub>771</sub>	1.4 ± 0.1	3.01 ± 0.68	40.6
human IgG	ND	3.52 ± 0.5	21 days <sup>e</sup>

<sup>a</sup> determined in humans

<sup>b</sup> KD was determined by the method of Anderson *et al.*, "J. Immunol." 125:2735-2741 (1980).

<sup>c</sup> determined by the method of Smith *et al.*, "Science" 238:1704-07 (1987).

<sup>d</sup> 5 residues 1-368 only

<sup>e</sup> + + The adhesion variant was injected intravenously into rabbits and samples of blood were collected periodically and assayed for the presence of the adhesion variant.

#### Claims

- 1 Nucleic acid encoding an amino sequence variant of an adhesion.
- 2 The nucleic acid of claim 1 wherein the adhesion is a CD4 polypeptide.
- 3 The nucleic acid of claim 2 wherein the variant is a CD4 polypeptide in which nucleic acid encoding the transmembrane domain has been modified whereby the CD4 polypeptide encoded thereby contains an inactivated transmembrane domain.
- 4 The nucleic acid of claim 3 wherein the transmembrane domain has been inactivated by its deletion or by substituting for the transmembrane domain an amino acid sequence having a substantially hydrophilic hydrophathy profile.

- 5 The nucleic acid of claim 2 wherein the variant comprises a fusion of (a) a polypeptide different from the CD4 and (b) a CD4 polypeptide
- 6 The nucleic acid of claim 5 wherein the polypeptide different from the CD4 bears a non-CD4 immune epitope
- 7 The nucleic acid of claim 6 wherein the polypeptide different from CD4 is fused to the amino or carboxyl terminus of mature CD4 and the transmembrane domain of CD4 has been inactivated.
- 8 The nucleic acid of claim 5 wherein the different polypeptide comprises a signal sequence.
- 9 The nucleic acid of claim 5 wherein the different polypeptide contains about from 5 to 300 residues.
- 10 The nucleic acid of claim 9 wherein the different polypeptide is capable of eliciting a humoral immune response in an animal
- 11 The nucleic acid of claim 10 wherein the different polypeptide is a viral polypeptide or an allergen
- 12 The nucleic acid of claim 1 wherein the different polypeptide is a human plasma protein having a plasma half life greater than from which the transmembrane domain has been deleted
- 13 The nucleic acid of claim 12 wherein the variant is a fusion of a polypeptide comprising at least one V-like domain of CD4 fused with a polypeptide comprising an immunoglobulin constant domain.
- 14 The nucleic acid of claim 1 wherein the adhesion is CD4, CD8 or the high affinity IgE receptor
- 15 The nucleic acid of claim 2 wherein the variant consists essentially of the V<sub>H</sub> through V<sub>H</sub> or V<sub>L</sub> through V<sub>L</sub> regions of the CD4 antigen
- 16 The nucleic acid of claim 2 which consists essentially of the CD4 insert of pCD4ΔN1a.
- 17 The nucleic acid of claim 12 wherein the different polypeptide is albumin, a lipoprotein or transferrin
- 18 The nucleic acid of claim 8 wherein the signal sequence is a bacterial signal sequence
- 19 The nucleic acid of claim 15 wherein the variant consists essentially of CD4 residues 1-368
- 20 The nucleic acid of claim 16 wherein the variant consists essentially of CD4 residues 1-180
- 21 The nucleic acid of claim 13 wherein the immunoglobulin constant domain is the constant domain of an IgG heavy chain
- 22 The nucleic acid of claim 5 wherein the different polypeptide is a cytotoxic polypeptide
- 23 The nucleic acid of claim 5 wherein the cytotoxic polypeptide is a diphtheria toxin A
- 24 A composition comprising an adhesion amino acid sequence variant which is incapable of cell membrane anchorage
- 25 The composition of claim 24 wherein the adhesion variant comprises a CD4 amino acid sequence capable of binding gp120
- 26 The composition of claim 25 further comprising an agent for inhibiting the aggregation of the variant selected from the group of a predetermined protein and a surfactant.
- 27 The composition of claim 26 wherein the agent is a surfactant.
- 28 The composition of claim 27 wherein the surfactant is Tween 80 or Tween 20.
- 29 The composition of claim 25 wherein the CD4 transmembrane domain has been deleted or has been substituted for by an amino acid sequence having a substantially hydrophilic hydrophobic profile.
- 30 The composition of claim 29 which is sterile and which further comprises a physiologically acceptable carrier
- 31 The composition of claim 25 wherein the variant comprises an immunoglobulin amino acid sequence
- 32 The composition of claim 31 wherein the immunoglobulin sequence comprises a constant domain sequence of an immunoglobulin heavy chain
- 33 The composition of claim 32 wherein the constant domain is linked at its N-terminus to the C-terminus of a transmembrane-deleted CD4 polypeptide.
- 34 The composition of claim 33 wherein the CD4 polypeptide contains V<sub>H</sub>V<sub>L</sub>
- 35 The composition of claim 33 wherein the CD4 polypeptide contains V<sub>H</sub>V<sub>L</sub>V<sub>H</sub>V<sub>L</sub>
- 36 The composition of claim 31 wherein the variant is in the form of a dimer
- 37 The composition of claim 36 wherein the composition comprises a fusion of a CD4 V-like domain to an immunoglobulin heavy chain constant domain
- 38 The composition of claim 24 wherein the variant is selected from the group consisting of:
  - (a) AC<sub>1</sub>;
  - (b) AC<sub>1</sub>-AC<sub>1</sub>;
  - (c) AC<sub>1</sub>-[AC<sub>1</sub>, AC<sub>1</sub>-AC<sub>1</sub>, AC<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>, V<sub>H</sub>C<sub>1</sub>-AC<sub>1</sub>, or V<sub>H</sub>C<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>];
  - (d) AC<sub>1</sub>-AC<sub>1</sub>-[AC<sub>1</sub>, AC<sub>1</sub>-AC<sub>1</sub>, AC<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>, V<sub>H</sub>C<sub>1</sub>-AC<sub>1</sub>, or V<sub>H</sub>C<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>];
  - (e) AC<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>-[AC<sub>1</sub>, AC<sub>1</sub>-AC<sub>1</sub>, AC<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>, V<sub>H</sub>C<sub>1</sub>-AC<sub>1</sub>, or V<sub>H</sub>C<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>];
  - (f) V<sub>H</sub>C<sub>1</sub>-AC<sub>1</sub>-[AC<sub>1</sub>, AC<sub>1</sub>-AC<sub>1</sub>, AC<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>, V<sub>H</sub>C<sub>1</sub>-AC<sub>1</sub>, or V<sub>H</sub>C<sub>1</sub>-V<sub>H</sub>C<sub>1</sub>]

ig) [A-Y]<sub>n</sub>-(V<sub>L</sub>C<sub>L</sub>-V<sub>H</sub>C<sub>H</sub>)

wherein A is a CD4 polypeptide containing a CD4 variable region-like domain; V<sub>L</sub>, V<sub>H</sub>, C<sub>L</sub> and C<sub>H</sub> represent light or heavy chain variable or constant domains of an immunoglobulin, n is an integer; and Y designates the residue of a covalent cross-linking agent

39 The composition of claim 38 wherein the V<sub>L</sub> and V<sub>H</sub> domains are capable of binding a predetermined antigen.

40 The composition of claim 31 wherein the immunoglobulin sequence is obtained from IgG1, IgG2, IgG3, IgG4, IgA, IgE, IgD or IgM

41 The composition of claim 25 wherein the variant comprises a polypeptide different from CD4 which is nonimmunogenic in humans.

42 The composition of claim 41 wherein the variant comprises a polypeptide which is immunogenic in humans.

43 The composition of claim 41 wherein the variant comprises a polypeptide having a human plasma half life which is greater than about 20 hours

44 The composition of claim 41 wherein the variant comprises a human transferrin, apolipoprotein or albumin polypeptide

45 The composition of claim 25 wherein the variant comprises a cytotoxic polypeptide

46 The composition of claim 45 wherein the cytotoxic polypeptide is ricin A chain or diphtheria toxin A.

47 A polypeptide comprising a CD4 amino acid sequence capable of binding gp120 which is cross-linked to (a) polypeptide having a plasma half life of greater than about 20 hours or (b) a cytotoxic polypeptide.

48 The polypeptide of claim 47 wherein the polypeptide of (a) is transferrin, an apolipoprotein or albumin.

49 The polypeptide of claim 47 wherein the cytotoxic polypeptide is cross-linked to the CD4 variable-like domain by a bifunctional cross-linking agent

50 A method for preparing an adhesion variant comprising transfecting a host cell with the nucleic acid of claim 1

51 A method for preparing an adhesion variant comprising recovering the variant from the culture of a host cell transfected with the nucleic acid of claim 1

52 The method of claim 51 wherein the adhesion is CD4 and the variant is recovered from the culture medium of the host cell

53 The method of claim 52 wherein the variant is recovered by adsorption onto a cation exchange resin.

54 The method of claim 53 wherein the variant is recovered by adsorption of contaminants onto an anion exchange resin.

55 The method of claim 52 wherein the variant lacks a functional transmembrane domain.

56 The method of claim 52 wherein wherein a salt is added to the culture medium to occupy charged domains of the variant, the resulting solution is contacted with a hydrophobic affinity chromatography resin to adsorb the variant, and the variant eluted from the resin by washing the resin with a decreasing gradient of salt.

57 The method of claim 52 wherein the variant is recovered by immunoaffinity chromatography.

58 The method of claim 57 wherein the immunoaffinity chromatography is directed against a polypeptide different from CD4 which is fused to CD4.

59 A method for the treatment of an HIV infection comprising administering to a patient infected with HIV a therapeutically effective dose of an amino acid sequence variant of CD4.

60 A replicable vector comprising the nucleic acid of claim 1

61 A process which comprises the preparation of nucleic acid according to any one of claims 1 to 23, a composition according to any one of claims 24 to 46 or a polypeptide according to any one of claims 47 to 49.

89131855



Fig. 1a

**1 6 7 0**

Fig. 1b

```

sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I
avall       avall       avall       avall       avall       avall       avall       avall       avall       avall
ppuMI       ppuMI       ppuMI       ppuMI       ppuMI       ppuMI       ppuMI       ppuMI       ppuMI       ppuMI
ecoO        ecoO        ecoO        ecoO        ecoO        ecoO        ecoO        ecoO        ecoO        ecoO
alwNI       alwNI       alwNI       alwNI       alwNI       alwNI       alwNI       alwNI       alwNI       alwNI
hinFI       hinFI       hinFI       hinFI       hinFI       hinFI       hinFI       hinFI       hinFI       hinFI
nlaIII      nlaIII      nlaIII      nlaIII      nlaIII      nlaIII      nlaIII      nlaIII      nlaIII      nlaIII
avall       avall       avall       avall       avall       avall       avall       avall       avall       avall
elul        elul        elul        elul        elul        elul        elul        elul        elul        elul
msel        msel        msel        msel        msel        msel        msel        msel        msel        msel
1201 CTCGGGACAGGTCCTGCTGGAATCCAACTCAAGGTTCTGCCCATGGTCCACCCAGCCTTTAATGCCGTAGCTTTATCAGACTTAATGCTAACCCA
GAGCCCTCTCCAGGACGACCTTAGGTTGTAGTTCCAAAGACGGGTGTACCAAGGCTGGGCTCGAAATTACGCATCAATAGTGTCAATTTAACGATTCCGT
350  SerGlyClnValLeuLeuGluSerAsnIleLeuValLeuProThrTrpSerThrProSerPheAsnAlaValValtyrHisSerOC.

nlaIV       nlaIV       nlaIV       nlaIV       nlaIV       nlaIV       nlaIV       nlaIV       nlaIV       nlaIV
banI        banI        banI        banI        banI        banI        banI        banI        banI        banI
fokI        fokI        fokI        fokI        fokI        fokI        fokI        fokI        fokI        fokI
hphI        hphI        hphI        hphI        hphI        hphI        hphI        hphI        hphI        hphI
scrFI       scrFI       scrFI       scrFI       scrFI       scrFI       scrFI       scrFI       scrFI       scrFI
bstNI       bstNI       bstNI       bstNI       bstNI       bstNI       bstNI       bstNI       bstNI       bstNI
sfanI       sfanI       sfanI       sfanI       sfanI       sfanI       sfanI       sfanI       sfanI       sfanI
rseI        rseI        rseI        rseI        rseI        rseI        rseI        rseI        rseI        rseI
mspI        mspI        mspI        mspI        mspI        mspI        mspI        mspI        mspI        mspI
hpaII       hpaII       hpaII       hpaII       hpaII       hpaII       hpaII       hpaII       hpaII       hpaII
1301 GTCAGGCACCGTGTATGAAATCTAACAATGCCGTCATCGTCATCTCGGACCGGTACCGCTGGATGCTGTAGGCATAGGCTTGCTTATGCCGGTACTGCC
CAGTCCGTGGCCACATACACTTTAGATTGTACGCCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGACCAATACGCCCATGACGG

nli         nli         nli         nli         nli         nli         nli         nli         nli         nli
haeIII      haeIII      haeIII      haeIII      haeIII      haeIII      haeIII      haeIII      haeIII      haeIII
sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I      sau96I
GGCCCTCTTCCGGGAT
CCCGGAGAACGCCCTA

```

89131855

1 1672  
 1 AGGTTTCAGCGGAAACACCACTACCCCGATCATCAGTTATTCCTTAAGGTCTCTTTTGTGTGGTTCGGTATGGGGGAGCTGCCCGCCAGGTTGG  
 TTCGAAGTCGGCTTGGTGGTGGCTAGTAGTCACTAGTCACTAGGANTCCAGAGAAACACACACCCAGGCCATACCCCGCTGACGGCGGTCCAAACC  
 MetGlyGlyThrAlaAlaArgLeuGly  
 10  
 101 GGGCGGTGATTTGTTTGTGTCATAGTGGCGCTCATGCGCTGCGGCAAAATATGCTTGGCGATGCTCTCTCAAGATGGCGACCCCAATCGATT  
 CCGGCACCTAAACACACGAGTATCACCCGGAGTACCCCGGCGCTTATACGGAACCCCTACGGAGAGTCTTACCGCTGGGTAGCTAA  
 10 AlaValIleLeuPheValValIleValGlyLeuHisGlyValArgGlyLysTyraLeuAlaAlaSerLeuLysMetAlaAspProAsnArgPhe  
 201 TCGCGGCAAAAGACCTTCGGTCTGACAGCTGCTGCGGAGGAAACAAAGTGTGTGGGCAAAAGGGGATACAGTGGAACTGACCTGTACAGCT  
 AGCCCGTTCTGGAAGCCAGGACCTGTGCGAGAGCTGCTCCCTTGTTCACACGACCCGTTTCCCTATGTCACCTGACATGACATGCGA  
 43 ArgGlyLysAspLeuProValLeuAspGlnLeuLeuGlnGlnGlyAsnLysValValLeuGlyLysGlyAspThrValGluLeuThrCysThrAla  
 301 TCCAGAGAAGAGCATACAATTCACCTGGGAAACTCCAAACAGATAAAGATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAAAGCTGA  
 AGGTCTCTCTGATGTTAAGGTGACCTTTTGAAGTGGTCTATTCTTAAGACCTTTAGCCCGAGGAAGAAATGATTTCCAGGTAGGTTCCACT  
 76 SerGlnLysLysSerIleGlnPheHisIleTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSerLysLeuAsn  
 401 ATGATCGCGCTGACTCAAGAGAAGCCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACCTACATCTGTGA  
 TACTAGCGGCACTGAGTCTTCTTCGGAACCCCTGTTCTTGAAGGGGACTAGTCTTCTAGAAATCTATCTCTGAGTCTATGAATGTAGACACT  
 110 AspArgAlaAspSerArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysAlaGluAspSerAspThrThrIleCysGlu  
 501 AGTGGAGGACCAAGGAGGAGGTGCATTTGCTAGTGTTCGGATTGACTGCCAATCTGTGACACCCACCTGCTTCAGGGGAGAGCTGACCTGACCTTG  
 TCACCTCTGCTCTCTCCACGCTTAACGATCACAGCTTAACGAGCTGAGCTGTGGTGGAGGAGTCCCGCTCTGGACTCGGACTCGGACTCGAAC  
 143 ValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThrLeu

Fig. 2a

89131855

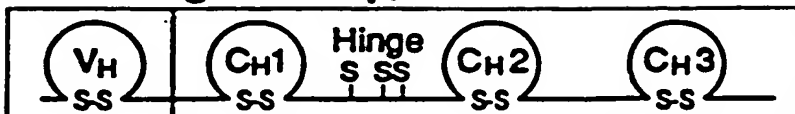
1 scrFI  
 2 bstXI  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630  
 631  
 632  
 633  
 634  
 635  
 636  
 637  
 638  
 639  
 640  
 641  
 642  
 643  
 644  
 645  
 646  
 647  
 648  
 649  
 650  
 651  
 652  
 653  
 654  
 655  
 656  
 657  
 658  
 659  
 660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678  
 679  
 680  
 681  
 682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703  
 704  
 705  
 706  
 707  
 708  
 709  
 710  
 711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725  
 726  
 727  
 728  
 729  
 730  
 731  
 732  
 733  
 734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754  
 755  
 756  
 757  
 758  
 759  
 760  
 761  
 762  
 763  
 764  
 765  
 766  
 767  
 768  
 769  
 770  
 771  
 772  
 773  
 774  
 775  
 776  
 777  
 778  
 779  
 780  
 781  
 782  
 783  
 784  
 785  
 786  
 787  
 788  
 789  
 790  
 791  
 792  
 793  
 794  
 795  
 796  
 797  
 798  
 799  
 800  
 801  
 802  
 803  
 804  
 805  
 806  
 807  
 808  
 809  
 810  
 811  
 812  
 813  
 814  
 815  
 816  
 817  
 818  
 819  
 820  
 821  
 822  
 823  
 824  
 825  
 826  
 827  
 828  
 829  
 830  
 831  
 832  
 833  
 834  
 835  
 836  
 837  
 838  
 839  
 840  
 841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850  
 851  
 852  
 853  
 854  
 855  
 856  
 857  
 858  
 859  
 860  
 861  
 862  
 863  
 864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876  
 877  
 878  
 879  
 880  
 881  
 882  
 883  
 884  
 885  
 886  
 887  
 888  
 889  
 890  
 891  
 892  
 893  
 894  
 895  
 896  
 897  
 898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916  
 917  
 918  
 919  
 920  
 921  
 922  
 923  
 924  
 925  
 926  
 927  
 928  
 929  
 930  
 931  
 932  
 933  
 934  
 935  
 936  
 937  
 938  
 939  
 940  
 941  
 942  
 943  
 944  
 945  
 946  
 947  
 948  
 949  
 950  
 951  
 952  
 953  
 954  
 955  
 956  
 957  
 958  
 959  
 960  
 961  
 962  
 963  
 964  
 965  
 966  
 967  
 968  
 969  
 970  
 971  
 972  
 973  
 974  
 975  
 976  
 977  
 978  
 979  
 980  
 981  
 982  
 983  
 984  
 985  
 986  
 987  
 988  
 989  
 990  
 991  
 992  
 993  
 994  
 995  
 996  
 997  
 998  
 999  
 1000

Fig. 2c

# CD4



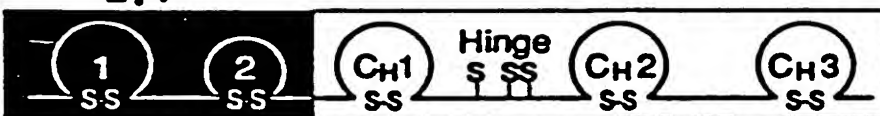
# Immunoglobulin $\gamma_1$



# Soluble rCD4



# CD4 $\gamma_1$



# CD4 $\gamma_1$



Fig.3

89131855

89131855



sau96I  
 avaiI  
 nlaIV  
 mnlI styI  
 1 GAATTCACCTCACCACATCAGCGCGCGGAGCCCTGACACATTTTGCACATTAATGAGAGTTTGTGCTCGTCACTTTGGGGGAGGACCA  
 CTTAAGTCAGAGTGGTGTAGTCGCGGAGCTCGGACTTCTAAACGTTGAATAGCTGTTGTCATTTCTCAACACGAGGAGTGAAGCCGCTCCCTGGT  
 72 ThrLeuThrIleSerGlyLeuGlnProGluAspPheAlaThrTyrCysGlnGlnTyrLysSerLeuSerLeuThrPheGlyGlyThrLys  
 sau3AI  
 dpnI  
 101 AGGTGGAGATCAACAGCACTCTGGTGGACCATCTGTCTCTCCGCCATCTGATGAGCAGTGAATCTGGAATCGCCTCTGTTGTGCTGCTGCT  
 TCCACCTTAGTTGCTTGACACCGAGCTGGTAGACAGAGTAGAAGGCGGTAGACTACTCGTCACTTTAGACCTTGACGGGAGACAACACACGACGA  
 104 ValGluIleLysArgThrValAlaAlaProSerValPheIlePheProSerAspGluGlnLeuLysSerGlyThrAlaSerValValCysLeuLeu  
 ← V<sub>6</sub>J<sub>2</sub> (C) →  
 haeIII  
 xmnI  
 201 GAATAACTTCTATCCAGAGAGAGGCCAAGTACAGTGGAGGTGGATAACGCCCTCCANTCGGGTAACCTCCACAGAGAGTGTCAACAGACGACGACGACGAAG  
 CTTATTGAAGATAGGGTCTCTCCGGTTCATGCTCCACCTATTGGGAGGTAGCCCATTTAGGGTCTCTCACAGTCTCTCGTCTGCTGCTTC  
 117 AsnAsnPheTyrProArgGluAlaLysValGlnTrpLysValAspAsnAlaLeuGlnSerGlyAsnSerGlnGluSerValThrGluGlnAspSerLys  
 haeIII  
 fnu4HI  
 ddeI  
 301 GACAGCACCTACAGCCTCAGCAGCAGCCTGACGCTGAGCAAGCAGACTACGAGAAACACAAAGTCTACGCCCTGCGAAGTCAACCATCAGGGCTGAGCT  
 CTGTCGTGGATGTCGGAGTCGTCTGGAGCTCGGACTCGGACTGCTTCTGCTGATGCTCTTTGTGTTTCAGATCGGACGCTTACGTGGGTAGTCCCGACTCGA  
 170 AspSerThrTyrSerLeuSerSerThrLeuThrLeuSerLysAlaAspTyrGluLysValTyrAlaCysGluValThrIleGlnGlyLeuSerSer  
 aluI  
 401 CGCCCGTCACAAAGCTTCAACAGGGGAGAGTGTAGAGGGAGAGTCCCCCACCCTGCTCCTCAGT  
 CGGGGAGTGTTCGAGAGTTGTCCCTCTCACATCTCCCTCTTCACGGGGGTGGACGAGGAGTCA  
 204 ProValThrLysSerPheAsnArgGlyGluCysAsn  
 ddeI  
 mnlI  
 bsp1286 bspMI mnlI  
 401 CGCCCGTCACAAAGCTTCAACAGGGGAGAGTGTAGAGGGAGAGTCCCCCACCCTGCTCCTCAGT  
 CGGGGAGTGTTCGAGAGTTGTCCCTCTCACATCTCCCTCTTCACGGGGGTGGACGAGGAGTCA  
 204 ProValThrLysSerPheAsnArgGlyGluCysAsn  
 haeIII  
 mnlI  
 bsp1286  
 haeIII  
 sau96I aluI  
 ecoO banII  
 hphI  
 301 GACAGCACCTACAGCCTCAGCAGCAGCCTGACGCTGAGCAAGCAGACTACGAGAAACACAAAGTCTACGCCCTGCGAAGTCAACCATCAGGGCTGAGCT  
 CTGTCGTGGATGTCGGAGTCGTCTGGAGCTCGGACTCGGACTGCTTCTGCTGATGCTCTTTGTGTTTCAGATCGGACGCTTACGTGGGTAGTCCCGACTCGA  
 170 AspSerThrTyrSerLeuSerSerThrLeuThrLeuSerLysAlaAspTyrGluLysValTyrAlaCysGluValThrIleGlnGlyLeuSerSer

Fig.5